Biophysical Studies on the RNA Cores of Satellite Tobacco Mosaic Virus

John Day, Yuri G. Kuznetsov, Steven B. Larson, Aaron Greenwood, and Alexander McPherson
Department of Molecular Biology and Biochemistry, University of California-Irvine, Irvine, California 92697-3900 USA

ABSTRACT Satellite tobacco mosaic virus (STMV) was probed using a variety of proteases. Consequences of the degradation were analyzed using gel electrophoresis, quasi-elastic light scattering (QELS), and atomic force microscopy (AFM). Proteolysis rates of 30 minutes for complete degradation of the protein capsid, up to many hours, were investigated. With each protease, degradation of virions 17 nm in diameter was shown by QELS to result in particles of 10 nm diameter, which is that of the RNA core observed in the virion by x-ray diffraction analysis. This was verified by direct visualization with atomic force microscopy. Using QELS, it was further shown that freshly prepared RNA cores remain as individual, stable, 10-nm condensed particles for 12 to 24 h. Clusters of particles then formed, followed by very large aggregates of 500 to 1000 nm diameter. AFM showed that the aggregates were composed of groups of the condensed RNA cores and were not due to unfolding of the nucleic acid. No unfolding of the core particles into extended conformation was seen by AFM until the samples were heated well beyond 90°C. Mass spectrometry of RNA core particles revealed the presence of a major polypeptide whose amino acid sequence corresponded to residues 2 through 25 of the coat protein. Amino acids 13 through 25 were previously observed to be in direct contact with the RNA and are presumably protected from protease digestion. Low resolution difference Fourier analyses indicated the courses of the remainders of the amino terminal strands (amino acids 2–12) in intact virions. Any individual strand appears to have several choices of path, which accounts for the observed disorder at high resolution. These positively charged strands, serving as virtual polyamines, engage the helical segments of RNA. The intimate association of amino acid residues 2 through 25 with RNA likely contributes to the stability of the condensed conformation of the nucleic acid cores.

INTRODUCTION

The structure of satellite tobacco mosaic virus (STMV; Valverde and Dodds, 1986, 1987) was solved and refined to a resolution of 1.8 Å using x-ray crystallographic methods (Larson et al., 1993a,b, 1998). This analysis revealed the presence of double helical RNA segments comprising about 45% of the entire genome of 1058 bases (Mirkov et al., 1989) lying just beneath the protein shell and in intimate contact with capsid dimers. The helical segments were presumed to be elements of stem-loop structures formed when the RNA was condensed and encapsidated by the protein. The identification of a free nucleotide, strongly associated with the protein but at some distance from the termini of the helical segments, suggested the likely dispositions of single-stranded loop elements. In total, the stem-(inferred) loop elements would account for about 80% of the total nucleic acid in the virion. The remainder of the RNA must connect, in some manner, the sequence of stem-loop structures.

As described previously (Larson et al., 1993a,b, 1998), extensive and intricate interactions exist between the RNA and capsid proteins. The involvement of protein with RNA is, however, probably even greater than revealed by the x-ray analysis, as residues 1 through 12 were not visualized in electron density maps. The amino terminal polypeptides, which disappear into the interior of the virion near the threefold axes, each bear six positive charges (two arginine, three lysine, and one amino terminal residue; Mirkov et al., 1989). They are likely involved in electrostatic interactions with the polyphosphate backbone of the RNA, and could alone account for neutralization of about 35% of the negative charge on the nucleic acid.

Details of the physical and chemical properties of the STMV RNA cores are of significant interest because with STMV it may be possible for the first time to deduce the complete conformation of the encapsidated nucleic acid for any virus. In addition, the interactions between the RNA and the protein are so extensive that some profound constraints on the mechanism of assembly seem evident. STMV may, therefore, serve as a model system for the encapsidation of genomic RNA in small viruses (Johnson and Rueckert, 1997).

In the experiments described here, using a variety of biophysical methods, we have attempted to define further the structural and functional relationship between the protein and nucleic acid components of the virus. The experiments utilized two strategies. One was based on low resolution difference Fourier maps to deduce the dispositions of the amino terminal strands of the protein subunits; the other degraded protein capsids using proteases while applying quasi-elastic light scattering (QELS), atomic force microscopy (AFM), gel electrophoresis, and mass spectrometry to evaluate the consequences.

MATERIALS AND METHODS

Proteolysis

To degrade STMV capsids, seven different proteases were investigated as well as cocktails composed of various combinations. These included bo-

Received for publication 23 October 2000 and in final form 6 February 2001.

Address reprint requests to Alexander McPherson, University of California-Irvine, Department of Molecular Biology and Biochemistry, Irvine, CA 92697-3900. Tel.: 949-824-1931; Fax: 949-824-1954; E-mail: amcphers@uci.edu.

0006-3495/01/05/2364/08 $2.00

vine and pig trypsin, subtilisin, nagarase, chymotrypsin, bovine pancreatic protease, and proteinase K. Digestions were carried out in Dulbecco's phosphate buffered saline (DPBS) at pH 7.2. There was little qualitative difference in the results obtained with the various proteases and protease combinations, though the rates at which proteolysis occurred varied substantially. The temperature for optimal proteolysis was also examined, and we found elevated temperatures to offer some advantage. Digestions were, therefore, performed at 65°C, whereas all other experiments and analyses were carried out, unless otherwise noted, at 25°C. For most of the proteolysis experiments, ribonuclease-free Proteinase K was used at a protease: virus ratio of 1:4. Under these conditions, degradation of the virus capsid was essentially complete within an hour. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to the procedures of Laemmli (1970) using mini-gels 10 cm × 10 cm in size and stained for protein using Coomassie blue.

Analysis of core particles

For preparation of the RNA core particles for gel electrophoresis and mass spectrometry, crystallized virus was further purified by chromatography on a 1 cm × 50 cm. Sephacryl S-100-HR column and concentrated to 4 mg/ml in Centricon concentrators (Millipore Corp., Bedford, MA). It was crystallized at 4°C by the addition of saturated ammonium sulfate until opalescence was observed in the solution and then allowed to stand for 4 to 10 days until crystallization was complete. The crystals were harvested by centrifugation and redissolved in water. We found it essential to crystallize STMV at least once and, if possible, multiple times before attempting any analysis of its RNA. This is generally done at 4°C and as rapidly as possible (4 to 10 days) so that a minimum of natural hydrolysis occurs. If it is not recrystallized, even after gel permeation chromatography, it remains significantly contaminated with fragments of tobacco mosaic virus (TMV). These especially confuse RNA gel analyses and related procedures with spurious TMV-RNA fragments.

To establish digestion conditions, purified and recrystallized STMV was subjected to a broad range of proteinase K concentrations, pH, temperature, and time regimens. The results were monitored using SDS-PAGE and, where useful, QELS. For preparation of RNA cores, 0.5 mg/ml virus was ultimately treated at 65°C with 0.125 mg/ml. Proteinase K for 40 min at pH 7.2. Degradation of the virus was monitored using QELS and was continued for an additional 10 min after 17-nm virions were no longer present. SDS-PAGE gels of the course of proteolysis of the virions show that even at 25°C virtually all of the coat protein is cleaved within minutes, and extensive digestion is complete within 5 to 10 minutes. The longer persistence of the 17-nm particles in the QELS spectrum during proteolysis suggests, however, that the extensively cleaved protein does not rapidly dissociate into fragments but maintains its structure for some time afterward.

RNA core particles were precipitated from the reaction mixture by addition of ethanol to 20% volume at 0°C. The precipitate was washed with DPBS and reprecipitated with ethanol twice more. RNA core particles were subjected to electrophoresis on denaturing, 2% formaldehyde agarose gels of 0.5 to 1.0%. Gels were stained with ethidium bromide and photographed under UV light. These show that the nucleic acid in the crystalline virus before digestion with proteinase K is completely intact. After proteolysis, some cleavage of the RNA is observed, presumably from contaminant enzymes, but it does not exceed about 3% of the RNA.

Mass spectrometry and amino acid sequencing

RNA cores prepared by proteinase K digestion were analyzed by MALDI-TOF mass spectrometry, using a PerSeptive Biosystems (Foster City, CA) Voyager DE-PRO Workstation with a matrix of sinapinic acid, to determine if polypeptides were associated with the RNA. A polypeptide was identified preliminarily by comparison of masses with those predicted from the known amino acid sequence. The amino acid sequence of the residual peptide from the RNA cores was then determined directly using a Hewlett-Packard (Palo Alto, CA) model 1000 A sequenator.

QELS

QELS measurements were carried out using a Malvern 4700c submicroparticle analyzer (Malvern Instruments, Inc., Southborough, MA) essentially as described for earlier QELS investigations of the crystallization of STMV (Malkin and McPherson, 1993, 1994; Malkin et al., 1993). Before any experiments were conducted, virus and protease solutions were centrifuged for 10 min at 12,000 × g and passed through 0.22-μm syringe filters to remove large foreign particles or aggregates. Samples analyzed by QELS had total volumes of 50 μl in square microcuvettes (Starna Cells, Inc., Atascadero, CA). The sample cell was maintained at the center of a glass cylinder containing temperature-controlled, double-distilled water that was continuously recirculated through a 0.22-μm filter.

The photomultiplier was positioned at 90° to the incident laser beam throughout the studies (Innova 70–3, 1 W argon laser, Coherent Laser Products Division, Palo Alto, CA). Measurements were made in photon correlation mode. The theory, technique, and methods of determination of particle size distribution from the analysis of the autocorrelation function have been described elsewhere (Berne and Pecora, 1976). In practice, this technique permits calculation of the translational diffusion coefficient of particles undergoing Brownian motion. If particles are assumed to have spherical shapes, which is true for STMV, the hydrodynamic radius r_h can be calculated using the Stokes-Einstein equation, which utilizes the diffusion coefficient D, together with the temperature T and viscosity η of the solvent, according to

$$r_h = kT/6\pi\eta D.$$

Sizes of aggregates were estimated from distributions of mass (i.e., the average was weighted according to relative mass of the scatterers). The system was calibrated for particle size using latex spheres of 14, 30, 40, 65, and 105 nm diameter (Interfacial Dynamics Co., Portland, OR).

AFM

For analysis by AFM (see Binning et al., 1986; Bustamante and Keller, 1995; Morris, 1994), samples of STMV or degraded STMV were adsorbed onto slips of freshly cleaved mica, which were then transferred into the sealed fluid cell of a Digital Nanoscope III atomic force microscope (Digital Instruments, Santa Barbara, CA). The entire volume of the cell, approximately 50 μl, was filled with buffer. Images were collected in both contact and tapping mode using oxide-sharpened silicon nitride tips from Digital Instruments. Most procedures were those described in earlier work on the crystallization of macromolecules as studied by AFM (Kuznetsov et al., 1997, 1999).

The measurement of particle size and the dimensions of features of individual particles must be approached with care. Sizes of individual particles adsorbed to the mica appear considerably larger because the image obtained is the convolution of the AFM tip shape with that of the particle. That is, the tip is not infinitely sharp and its curved surface immediately adjacent to the absolute tip causes vertical displacement of the cantilever and, therefore, gives rise to edges in the image before as well as after the absolute tip encounters the object. This does not, however, affect the total vertical displacement of the cantilever. As a consequence, single objects visualized by AFM appear broader than their true dimensions, but yield an accurate and precise vertical dimension. For roughly spherical particles, such as icosahedral viruses, although their width appears greater than is in fact the case, the vertical dimension (height) of the particles gives a remarkably accurate value for their true diameter, accurate to within 1 nm.

In the course of our studies on viruses, we have examined eight different varieties whose capsid sizes were accurately known. Our images of these viral particles consistently showed them to have lateral dimensions of about 2.5 times their actual diameter. This was confirmed by measuring center-to-center distances of particles in the crystalline lattices of the viruses when available, and by comparison with the detailed structures known from x-ray diffraction, electron microscopy, or light scattering. Therefore, particle diameters based on AFM cited here are based only on vertical measures of their diameters.

X-ray diffraction

Using x-ray diffraction data, programs, and procedures described previously for the 1.8 Å structure determination and refinement (Larson et al., 1998), 40 to 55 Å resolution Fo-Fc difference Fourier maps were computed. These were displayed using the program "O" (Jones and Kjeldgaard, 1994; Jones, 1978) on a 340 VGX Silicon Graphics computer, and as transparencies affixed to plastic sheets (minimaps). Superimposed on the difference electron density maps were the regions occupied by protein and that portion of the RNA visualized in the 1.8 Å Fourier maps. An entire quadrant of the virus was displayed for examination. Probable paths for the amino terminal strands beginning at amino acid residue 13 were first traced on the Plexiglas sheets as pseudo-Cα positions. Coordinates, measured by hand from the sheets and the paths, serving as linear sets of fiducial points, were then used to construct, using "O," polyalanine chains representing putative amino terminal strand dispositions. These were then examined for stereochemical reality and bad contacts, and further adjusted where appropriate. No attempt was made at these low resolutions to identify individual residues or specific interactions with the RNA or to refine the polyalanine backbones.

RESULTS

In parallel with observation by QELS and AFM of the effects of proteolysis on virion structure, the process was monitored by gel electrophoresis. When intact STMV particles were subjected to SDS-PAGE, only a single band at 15.5 kd, corresponding to the intact coat protein, appeared. Cleavage of virtually all coat protein occurred within 1 min of protease exposure to give a pattern of discrete fragments, which in ensuing minutes tended to lower and lower molecular weights. Within 30 min after STMV exposure to proteases the discrete capsid subunit band dissolved completely into a smear comprised of heterogeneous, lower molecular weight species. The integrity of the RNA following proteolytic digestion was confirmed by electrophoresis on agarose gels as described in Materials and Methods.

Solutions of STMV, when examined by QELS as shown in Fig. 1 *a*, yield distinctive, symmetrical peaks at 17 nm diameter, corresponding to the $T = 1$ icosahedral particles (Caspar and Klug, 1962; Dodds, 1991), though occasionally with a small shoulder corresponding to the dimer species (Malkin and McPherson, 1993; Malkin et al., 1993). In phosphate buffered saline at 22°C, the monomer peak is stable indefinitely. Addition of individual proteases or protease cocktails causes the peak, as shown in Fig. 1 *b*, to shift transiently to a particle diameter of about 20 to 25 nm, and then to decrease steadily to a position corresponding to a

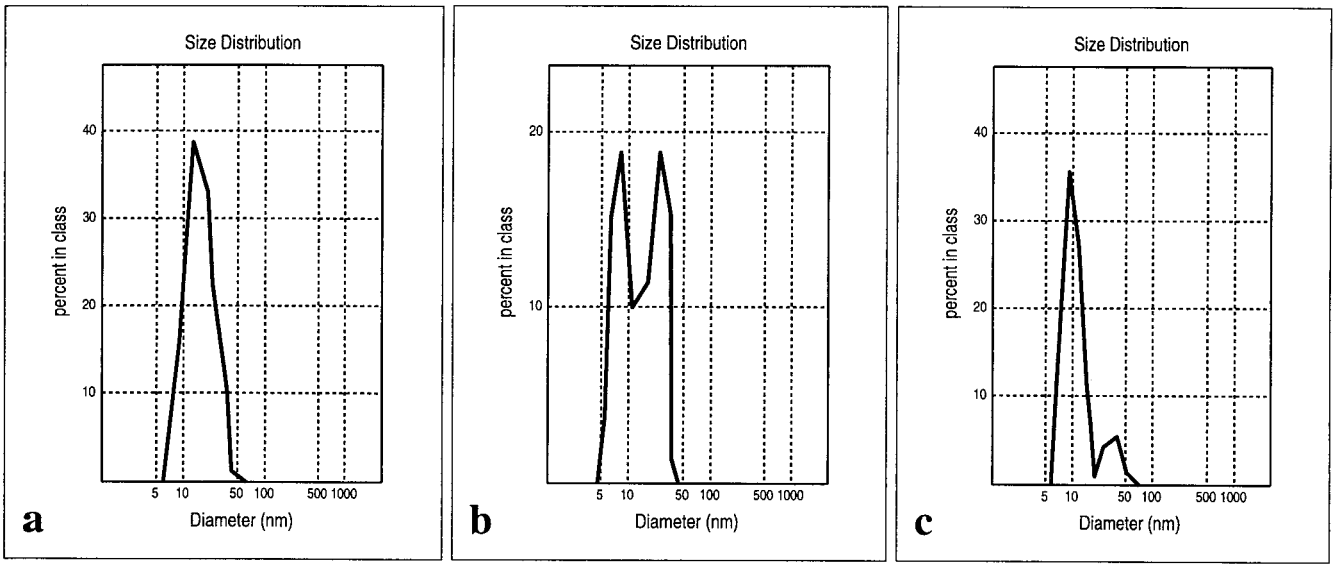

FIGURE 1 Distribution of particle sizes as measured by quasi-elastic light scattering. (*a*) Recrystallized STMV before addition of any protease shows a single peak corresponding to the diameter of the virus particle centered at 17 nm, but inclined slightly toward higher values due to the presence of some dimer species. After addition of a protease, in this case proteinase K, after about 15 to 30 min (*b*), the 17-nm peak at right has declined to about half its previous value and has shifted to the right, nearer to 23 nm diameter, and a peak corresponding to a particle size of about 10 nm appears to the left. After 30 min (*c*), the peak at 17 nm corresponding to intact particles has disappeared, and virtually all particles have assumed a 10-nm diameter. A second, broad peak representing clusters is also prominent at about 45 nm diameter. Ultimately the entire spectrum is obscured by very large aggregates >1000 nm in diameter.

particle diameter of about 10 nm. The value of 10 nm corresponds closely to the diameter of the RNA core, which is characterized, in part, by the icosahedrally distributed array of double helical RNA segments (Larson et al., 1993a,b).

Monitored by QELS, the RNA core particles of 10 nm diameter do not change in size nor aggregate significantly, but remain stable for about 24 h at 25°C in physiological saline. The initial, but transient increase in particle size from 17 nm to 20–25 nm we interpret to be a consequence of the binding of protease molecules to the intact virus, which would increase their apparent hydrodynamic radius. The transition of 20–25 nm particles to 10 nm particles tracks the degradation of the protein capsid.

The possibility of the 20–25 nm diameter peaks representing a swollen form of the virus is unlikely because the volume increase, were that the case, would be two- to threefold. This is insupportable. There is, in addition, no reason to suspect protease attack should inspire such a swelling, particularly of this magnitude. STMV does in fact swell as a consequence of high pH, but the diameter changes only from 17 nm to 18 nm (unpublished data).

Upon standing for about 24 h after formation of the 10 nm particles by proteolysis, another peak appears at about 45 nm diameter. This rises to represent several percent of the particles, as shown in Fig. 1 *c*. If samples are left undisturbed for more than 24 h at room temperature, peaks also begin to appear, corresponding first to species of about 500 nm, presumably large aggregates; these condense further with time to particles of 1000 nm or more in diameter.

Once the very large (>1000 nm) particles appear, even at a fraction of a percent of the total particles, their corresponding signal virtually precludes appearance in the spectrum of peaks resulting from smaller species still in solution. The presence of large particles, as were produced by aggregation of the RNA, masks the presence of smaller particles unless the aggregates are removed by centrifugation and filtration. This is a technical limitation of QELS, the spectrum of which is strongly dominated by large particle sizes. The very large aggregates (>500 nm) can be removed from the samples by 15 min of high-speed centrifugation followed by microfiltration. This restores the small particle diameters to the spectrum. Nonetheless, over sufficient time or upon exposure to higher temperatures, all of the material in the sample ultimately accumulates in the very large aggregates.

To investigate the stability of RNA cores resulting from capsid proteolysis, RNA samples were produced as above and then exposed to increasing temperature. Change in core size and conformation was again monitored using QELS. No change in the QELS spectrum was evident until a temperature of nearly 75°C was reached. At that temperature the peak of 10 nm declined, the broad peak centered at 45 nm appeared, and this was followed by peaks corresponding to sizes indicative of particles >1000 nm diameter. Essentially, the same progression of peaks was obtained with heat as with time.

To correlate particle sizes inferred from QELS with specific structures, AFM was utilized. Samples were of three types: (i) freshly purified STMV, which had been recrystallized two to four times (these exhibited a single peak at 17 nm by QELS and were composed only of intact STMV virions), (ii) fresh solutions which had been treated with proteases as described above, and (iii) protease-produced core particles more than 24 h old. Fig. 2 *a* is an AFM image of STMV virions adsorbed directly on a mica substrate. The heights of the virions as measured by AFM are 17 nm as expected from the x-ray determined structure. Fig. 2 *b* is an AFM image of STMV which have been degraded by proteases to core particles, as verified by QELS and SDS-PAGE. The particles, which fill the field, have globular appearances and heights of 10 nm, consistent with the

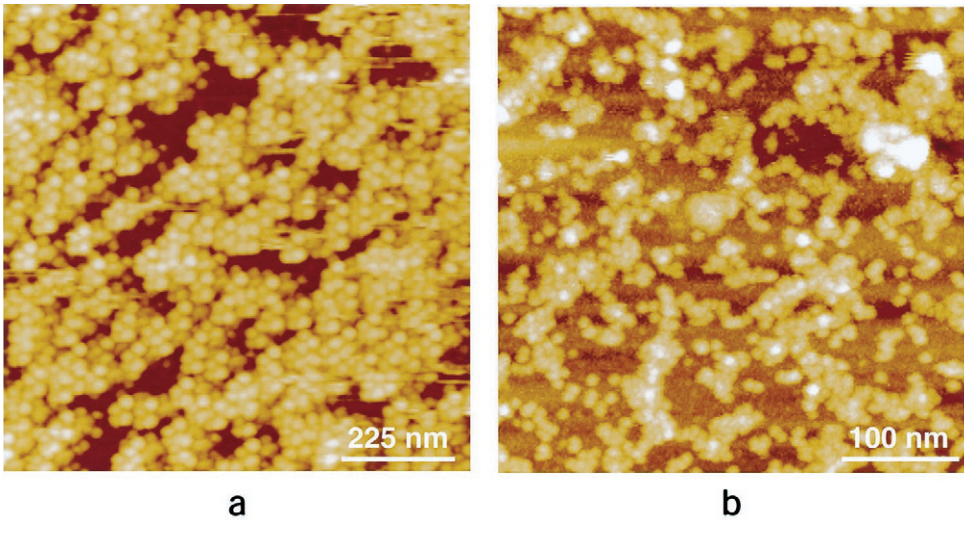

FIGURE 2 (*a*) STMV virions of 17 nm diameter are seen adsorbed to a mica surface. The individual spherical particles are clearly visible. After degradation with proteases, the resulting RNA core particles of 10 nm (from QELS) are seen adsorbed to the mica surface (*b*). Although aggregates predominate, undoubtedly promoted by adsorption to the mica, some individual RNA core particles are clearly visible. Both images were obtained by contact mode AFM in the fluid cell filled with water. Scan areas: 1 μm^2 (*a*) and 380 nm^2 (*b*).

condensed nucleic acid from the x-ray diffraction study (Larson et al., 1993b, 1998).

In Fig. 3 some of the core particles are seen at higher magnification using AFM, again, adsorbed onto mica in the fluid cell. Both individual particles and aggregates are visible, but particles incorporated into clusters seem to offer more detailed images, possibly because of greater stability of the larger masses and firmer attachment to the substrate. In Fig. 3, *b* and *c*, it can be seen that the core particles, from height measurements 10 nm in diameter, exhibit a substructure of puffs or swells that protrude slightly from the surface. Presumably these reflect the structure of the RNA which makes up the core particles. From inspection and from surface area considerations, we can estimate the number of the puffs, which are about 3 nm wide, to be about 30. The common features of the particles suggests that the RNA assumes a more or less uniform conformation when encapsidated that exhibits a fairly regular number and distribution of units on its outside surface. This is consistent with the symmetrical distribution of RNA stem-loops observed in the x-ray structure.

Mass spectrometry of RNA core particles was carried out with the objective of determining if portions of the coat protein, particularly the amino terminal polypeptides, remained associated with the RNA after proteolysis. This revealed that a peptide consistent with amino acids 2 through 25 was indeed present in the RNA cores, though quantitation by mass spectrometry of the amount of the peptide was not possible. Subsequent amino acid sequencing confirmed the peptide to be amino acids 2 through 25. The absence of the terminal methionine appears to be complete, and we suspect it may have been removed by host cell enzymes before encapsidation. Amino acids 2 through 12 were not previously visible in electron density maps, but amino acid residues 13 through 26 were closely involved in nucleic acid interactions along the inside surface of the virion (Larson et al., 1998). These could very well have been protected from protease digestion.

To investigate further the role of amino terminal strands in the organization and stability of the RNA cores, additional analyses based on the x-ray diffraction data were undertaken. An approach which proved informative was calculation of difference Fourier syntheses using as coefficients the differences between observed structure amplitudes, Fo, and structure amplitudes calculated from the 1.8 Å resolution-refined structure, Fc, but restricted to the resolution range ∞ to 5.0 Å. Phases were those calculated from the native virus structure (including protein, water, and RNA) over that same resolution range. The rationale for confining the calculation to this resolution range was that, although amino terminal strand dispositions might be disordered at a detailed level, at lower resolution, multiple, discrete paths might become evident and general motifs might emerge. A similar approach to defining disordered regions of virus structures, using low resolution data, was applied by Tsuruta et al. (1998).

Icosahedrally averaged, low resolution difference Fourier maps were characterized by a virtual void from the center of the virion to a radius of 35 Å. Between about 40 and 50 Å radius, however, the maps contained strong chains of continuous density, like those shown in Fig. 4, that entwined with the RNA helical segments. The nucleic acid helices lie just beneath the protein shell, the interior radius of which is about 55 Å. The bold chains of density occasionally interpenetrate one another or form foci where they converge and then separate. For the most part, the chains of density closely approach the volumes occupied by triad related RNA helices, on their most interior surfaces.

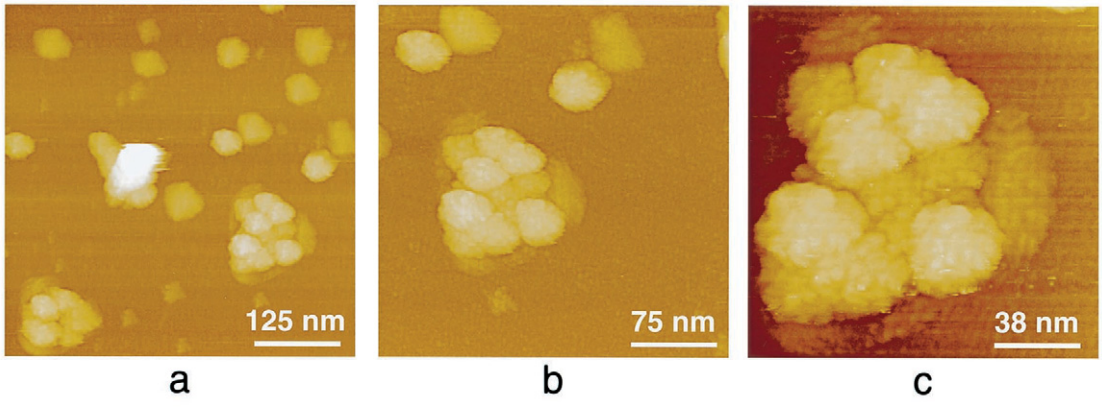

FIGURE 3 STMV virions after digestion with proteinase K and exhibiting diameters of 10 nm according to QELS are imaged here at higher magnification using atomic force microscopy. (*a*) Individual core particles as well as clusters of particles are seen adsorbed to freshly cleaved mica in the digestion buffer. Based on height measurement, the core particles have an average diameter of 10 nm, consistent with the value yielded by QELS. At higher magnification, (*b* and *c*), some substructure of the core particles, presumably reflecting the distribution and conformation of the encapsidated RNA, begins to emerge. The swells or puffs that cover the surface of the particle are about 3 nm in width. Tapping mode was used for this series of images. Scan areas: 500 nm^2 (*a*), 300 nm^2 (*b*), and 150 nm^2 (*c*).

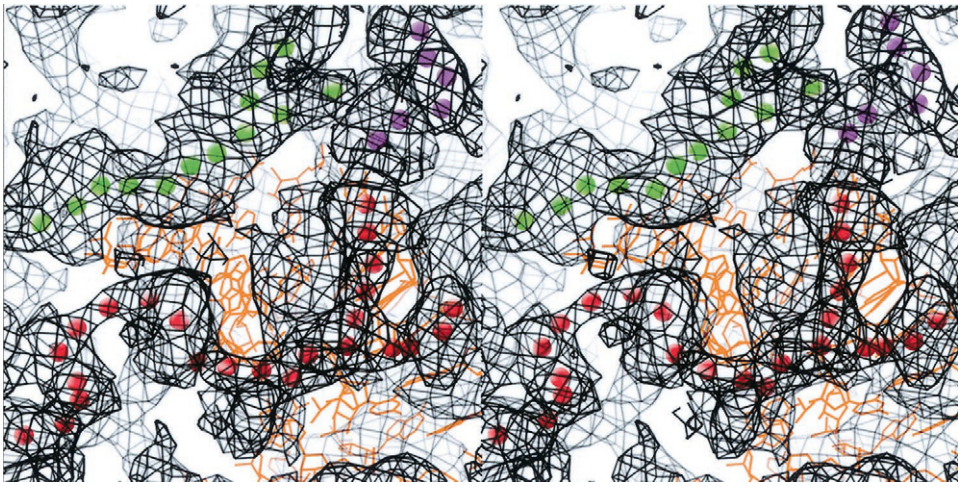

FIGURE 4 In this stereodiagram, a small portion of the low resolution difference Fourier map is superimposed on local RNA structure. Here, three continuous chains of density arising from amino terminal strands of coat protein (*red, green, and blue dotted paths*) are seen to embrace and interact closely with a helical segment of the nucleic acid, shown in yellow. The chains of density observed here are typical of density seen entwined with the RNA helices throughout the map.

As suggested by the position of the last amino acid visible in the 1.8 Å electron density map (threonine 13), the amino terminal strands from threefold related subunits converge near triad axes and, from those points, appear to penetrate the core of the nucleic acid. Because icosahedral particles, containing in their interiors asymmetrically distributed RNA, crystallize in 60 (capsid) equivalent orientations, electron density maps represent averages of the amino terminal strands distributed over the entire icosahedral particle. A volume in the electron density map common to many strands, or shared by strands otherwise having different dispositions, will appear as a nexus where strands coalesce. That is, one may observe several strands of density entering a central mass of electron density and several density strands emerging from it. Pairing entering strands with emerging strands in combination provides indications of the most populated courses followed by the various amino terminal polypeptides as they enter the RNA core.

Inspection of the difference electron density maps indicated that, from any one of the symmetry-related threonine 13 positions at a threefold axis, there are five popular choices available to an individual amino terminal strand, amino acids 2 through 12. These are shown in Fig. 5. It is important that at any threefold position only one of the possible paths will be utilized by an amino terminal polypeptide. One of these paths, however, originates at each of the three amino acid 13 start points, and, presumably (but not necessarily), one leads to each of the RNA helical segments. Because there are 60 protein subunits but only 30 RNA helices, on average at least, and perhaps in every case, there will be two amino terminal strands engaging every nucleic acid stem. The pattern seen in Fig. 5 is consistent with the difference electron density seen in the icosahedrally averaged low resolution difference Fourier synthesis.

There may be additional variations, particularly at unusual places in the capsid, for example, where the termini of the RNA molecules associate with the interior of the capsid. Furthermore, we see from the electron density maps that some of the paths may finish with the final few residues (amino acids 2–4) in different orientations as they engage the RNA helices. These are manifested as forks near the amino termini. These observations suggest that, aside from

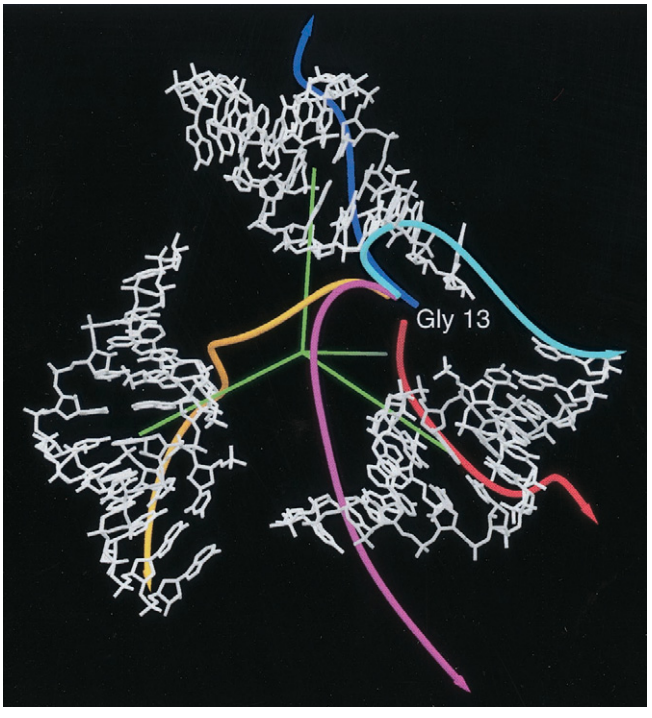

FIGURE 5 Diagram showing most popular paths assumed by the amino terminal peptides of each coat protein molecule. Each strand, bearing six positive charges, assumes predominantly one of the five courses indicated here, and interacts closely on the interior of the RNA core with one of three helical RNA segments related by icosahedral triad axes. Minor paths may also be available to the strands, and some among the five shown here are probably preferred over others. Three amino terminal strands make choices among the paths in the neighborhood of each icosahedral threefold axis. The position of the threefold axis is indicated, as are three perpendiculars to the icosahedral edges. Another line lies along a virion radius and is directed into the interior.

the promiscuous association with RNA helices, the residues of the amino terminal strands may exhibit some freedom in choosing the phosphate groups with which they associate, and how they position themselves with respect to those groups. That is, the amino terminal strands exhibit a good deal of structural fluidity.

DISCUSSION

Although QELS indicates that most of the capsid protein is degraded and lost upon proteolytic digestion, it does not imply that entire polypeptide chains are necessarily removed. Indeed, in intact virions, amino terminal strands penetrate into the interior, where their seven positive charges interact with the ribose phosphate backbone of the nucleic acid. They are apparently not released, even after proteolysis of the remainder of the protein subunits, as shown by mass spectrometry analysis of the core particles.

The QELS results imply that the RNA cores, prepared by proteolytic removal of the capsid protein but still associated with positively charged amino terminal polypeptides, maintain a stable, condensed conformation as 10-nm-diameter particles, not substantially different than in the intact virion. The protein shell is undoubtedly complicit in folding the RNA into its encapsidated form, but, except for the amino terminal strands, intact protein subunits do not appear necessary to maintain the conformation. The stability of the RNA cores is considerable, as they show no evidence of disruption by AFM imaging until the temperature exceeds about 90°C. The large peak in the QELS spectrum that appeared at 75°C was, in fact, due to aggregation of the core particles, and this was also apparent from AFM. Amino terminal strands, carrying a total of 360 positive charges (60×6), apparently serve as extended polyamines that promote cohesion within core particles. AFM images suggest that, consistent with x-ray results, the RNA exhibits a more or less regular substructure on its surface.

Low resolution difference Fourier maps based on the refined, 1.8 Å resolution structure, which contains 10,080 water molecules as well as capsid proteins (excepting amino acids 2–12) and about 45% of the genomic RNA (Larson et al., 1998), reveal the most probable paths of the amino terminal strands, amino acids 2 through 12. These paths suggest a fluid character which, along with their diversity, explains the disorder they exhibit and their invisibility in high resolution Fourier maps. The fluid nature of these terminal strands, coupled with their high positive charge density, produces a means of neutralizing the negative phosphate groups of the RNA helical segments. The intricacy of their arrangement at the threefold axes and their intimate association with the RNA helices further suggests that they provide not only structural maintenance, but may act as organizing agents as well. It seems likely that they are the protein subunit elements to first engage the RNA, through electrostatic interactions.

Whereas the amino terminal strands may be crucial in engaging the nucleic acid and maintaining its encapsidated conformation, the remainder of the protein subunits seem incidental in this regard. The β barrels of the protein subunits do, however, carry out an organizing function and are probably responsible, through protein-protein interactions, for drawing the entire RNA-protein aggregate together into an ordered icosahedral virion. The amino terminal strands stabilize the helix-loop elements of the RNA and provide an electrostatic adhesive that maintains the RNA in a defined conformation, whereas the remainder of the subunit serves to direct nucleic acid condensation into the icosahedral virion. The relationship seen here, both structural and functional, of the RNA helices and amino terminal strands is similar to that suggested by Argos (1981; see also Casjens, 1985; 1997) for amino terminal strand binding to helical RNA.

This research was supported by grants and contracts from the National Institutes of Health and the National Aeronautics and Space Administration.

REFERENCES

Argos, P. 1981. Secondary structure prediction of plant virus coat proteins. *Virology* 110:55–62.

Berne, B. J., and R. Pecora. 1976. *Dynamic Light Scattering*, John Wiley, New York.

Binning, G., C. F. Quate, and C. Gerber. 1986. Atomic force microscopy. *Phys. Rev. Lett.* 56:930–933.

Bustamante, C., and D. Keller. 1995. Scanning force microscopy in biology. *Phys. Today.* 48:32–38.

Casjens, S. 1985. Nucleic acid packaging by viruses. *In* Virus Structure and Assembly, ch. 3. Jones and Bartlett, Inc., Boston. 76–147.

Casjens, S. 1997. Principles of virion structure, function and assembly. *In* Structural Biology of Viruses. W. Chin, R. M. Burnett, and R. Garcea, eds. Oxford University Press, Oxford. 3–37.

Casper, D. L., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* 27:1–24.

Dodds, J. A. 1991. Structure and function of the genome of satellite tobacco mosaic virus. *Canad. J. Plant Pathol.* 13:192–195.

Johnson, J. E., and R. R. Rueckert. 1997. Packaging and release of the viral genome. *In* Structural Biology of Viruses. W. Chin, R. M. Burnett, and R. Garcea, eds. Oxford University Press, Oxford. 269–287.

Jones, T. A. 1978. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 11:268–272.

Jones, T. A., and M. Kjeldgaard. 1994. O: The Manual, Uppsala University Press, Uppsala, Sweden.

Kuznetsov, Y. G., A. J. Malkin, and A. McPherson. 1999. AFM studies of the nucleation and growth mechanisms of macromolecular crystals. *J. Cryst. Growth.* 196:489–502.

Kuznetsov, Yu. G., A. J. Malkin, T. A. Land, J. J. DeYoreo, A. P. Barba, and A. McPherson. 1997. Molecular resolution imaging of macromolecular crystals by atomic force microscopy. *Biophys. J.* 72:2357–2364.

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature.* 227:680.

Larson, S. B., S. Koszelak, J. Day, A. Greenwood, J. A. Dodds, and A. McPherson. 1993a. Double helical RNA in satellite tobacco mosaic virus. *Nature.* 361:179–182.

Larson, S. B., S. Koszelak, J. Day, A. Greenwood, J. A. Dodds, and A. McPherson. 1993b. Three-dimensional structure of satellite tobacco mosaic virus at 2.9 Å resolution. *J. Mol. Biol.* 231:375–391.

Larson, S. B., J. Day, A. Greenwood, and A. McPherson. 1998. Refined structure of satellite tobacco mosaic virus at 1.8 Å resolution. *J. Mol. Biol.* 277:37–59.

Malkin, A. J., J. Cheung, and A. McPherson. 1993. Crystallization of satellite tobacco mosaic virus. I. Nucleation phenomena. *J. Cryst. Growth.* 126:544–554.

Malkin, A. J., and A. McPherson. 1993. Crystallization of satellite tobacco mosaic virus. II. Postnucleation events. *J. Cryst. Growth.* 126:555–564.

Malkin, A., and A. McPherson. 1994. Light scattering investigations of nucleation processes and kinetics of crystallization in macromolecular systems. *Acta Cryst. D.* 50:385–395.

Mirkov, T. E., D. M. Mathews, D. H. DuPlessis, and J. A. Dodds. 1989. Nucleotide sequence and translation of STMV-RNA. *Virology.* 170: 139–146.

Morris, V. J. 1994. Biological applications of scanning probe microscopies. *Prog. Biophys. Mol. Biol.* 61:31–185.

Tsuruta, H., V. S. Reddy, W. R. Wikoff, and J. E. Johnson. 1998. Imaging RNA and dynamic protein segments with low-resolution virus crystallography: experimental design, data processing and implications of electron density maps. *J. Mol. Biol.* 284:1439–1452.

Valverde, R. A., and J. A. Dodds. 1986. Evidence for a satellite RNA associated naturally with the U5 strain and experimentally with the U1 strain of TMV. *J. Gen. Virol.* 67:1875–1884.

Valverde, R. A., and J. A. Dodds. 1987. Some properties of isometric virus particles which contain the satellite RNA of TMV. *J. Gen. Virol.* 68:965–972.